Time Series Model in Frequency Domain

林昕宏[1]　徐捷耀[2]

[1]Department of Quantitative Finance
[2]Department of Computational and Modeling Science

June 16, 2023

Table of Contents

Table of Contents

Dataset

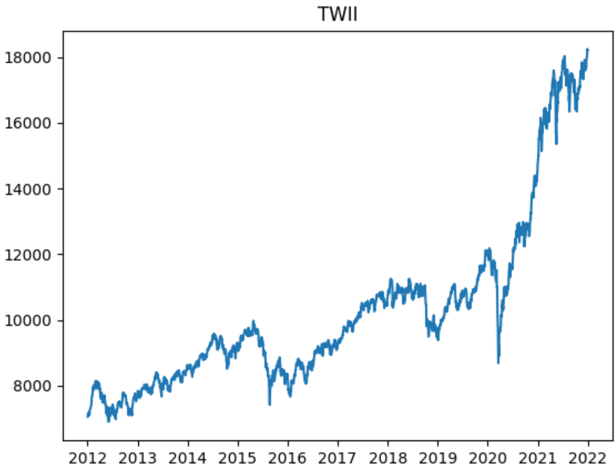


Figure: TWII close price from 2012 to 2022

Table of Contents

If a set of sequences can be expressed as $x[n], n = 0, 1, 2, \cdots, N-1$,its discrete Fourier transform (DFT) pair can be expressed as：

$$X(k) = \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi}{N}nk}, \tag{1}$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{i\frac{2\pi}{N}nk}. \tag{2}$$

$X(k)$: Fourier Coefficient（傅立葉係數）

Discrete Fourier transform

If a set of sequences can be expressed as $x[n], n = 0, 1, 2, \cdots, N-1$,its discrete Fourier transform (DFT) pair can be expressed as：

$$X(k) = \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi}{N}nk}, \tag{1}$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{i\frac{2\pi}{N}nk}. \tag{2}$$

$X(k)$: Fourier Coefficient（傅立葉係數）
So that DFT(1) can transform the $x[n]$ from 1-dimension to N-dimension.

Data DFT

For a given length m (usually choose the length of a week, half month, or a month), do DFT for every m data.

Figure: Data DFT for every m period

Transformed Data

We choose $m = 20$ and do DFT for data(length $= 2441$) every 20. Finally, the dimension of transformed data X_k is (122×20).

	0	1	2	...	18	19
0	146840+0j	690+2851j	-757+1074j	...	-757-1074j	690-2851j
1	159344+0j	-420+650j	-116-330j	...	-116+330j	-420-650j
...	...	...	...	...	...	...
120	351091+0j	-2281+170j	169-373j	...	169+373j	-2281-170j
121	356792+0j	1100+934j	-450+946j	...	-450-946j	1100-934j

Table: The part of transformed data X_k

Since X_k is conjugate symmetry, we only keep half of the data as X_k^* e.g. X_k^* is selected from the $1st$ to $11th$ column of X_k. The dimension of X_k^* is (122×11)

Transformed Data divide

Since the value in transformed data are complex numbers, they should be divided into real part and imaginary part.
Divide the X_k^* into the real part and imaginary part, and both of them have the same dimension (122×11).

Table of Contents

ADF test

	real	imag
idx_0	1.202	nan
idx_1	-5.942	-6.663
idx_2	-5.967	-10.61
idx_3	-10.805	-13.578
idx_4	-4.768	-2.351
idx_5	-7.189	-11.287
idx_6	-12.27	-9.409
idx_7	-13.422	-13.598
idx_8	-10.696	-6.263
idx_9	-9.965	-11.691
idx_10	-7.232	-11.012

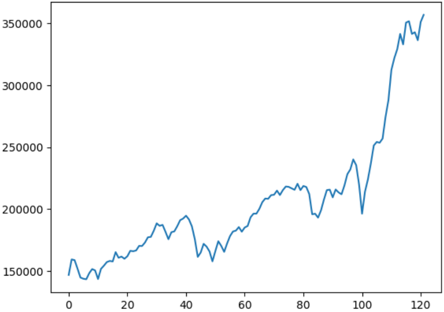
	Critical Values
10%	-2.58
5%	-2.89
1%	-3.49

We find that the first column of real part is not stationary.

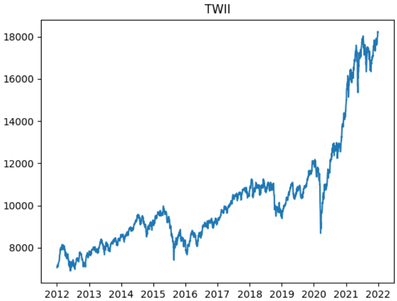
The test-statistic of every column of real part and imaginary part

Table of Contents

The pattern of first column in real part



The first column in real part



TWII close price from 2012 to 2022

Model Building

Since the first column of X_k^* has different properties to others, we divide X_k^* into three parts for three different models.

- part1: The first column of the real part in X_k^*.
- part2: The remaining columns of the real part in X_k^*.
- part3: The imaginary part except for the first column in X_k^*.

ACF

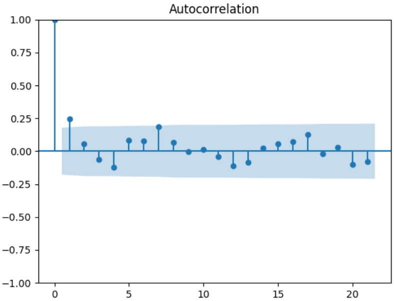


Figure: ACF of first-order difference

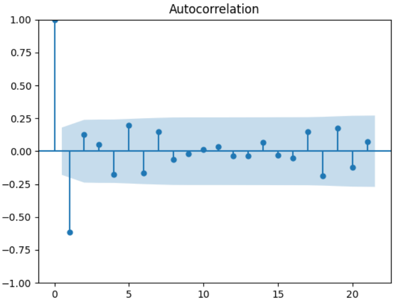


Figure: ACF of third order difference

Part1 model: ARIMA (AIC base)

Dep. Variable:	y	**No. Observations:**	122
Model:	SARIMAX(2, 3, 3)	**Log Likelihood**	-4039.661
Date:	Thu, 08 Jun 2023	**AIC**	8093.322
Time:	08:18:21	**BIC**	8112.950
Sample:	0 - 122	**HQIC**	8101.294
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	-405.8270	1432.492	-0.283	0.777	-3213.460	2401.806
ar.L1	-1.6154	0.029	-55.997	0.000	-1.672	-1.559
ar.L2	-0.8159	0.028	-29.223	0.000	-0.871	-0.761
ma.L1	1.0697	0.066	16.259	0.000	0.941	1.199
ma.L2	-0.4546	0.105	-4.330	0.000	-0.660	-0.249
ma.L3	-0.7705	0.075	-10.258	0.000	-0.918	-0.623
sigma2	2.273e+08	0.114	2e+09	0.000	2.27e+08	2.27e+08

Ljung-Box (L1) (Q):	1.25	**Jarque-Bera (JB):**	65842.49
Prob(Q):	0.26	**Prob(JB):**	0.00
Heteroskedasticity (H):	0.01	**Skew:**	10.54
Prob(H) (two-sided):	0.00	**Kurtosis:**	114.84

ARIMA(2,3,3): Residual

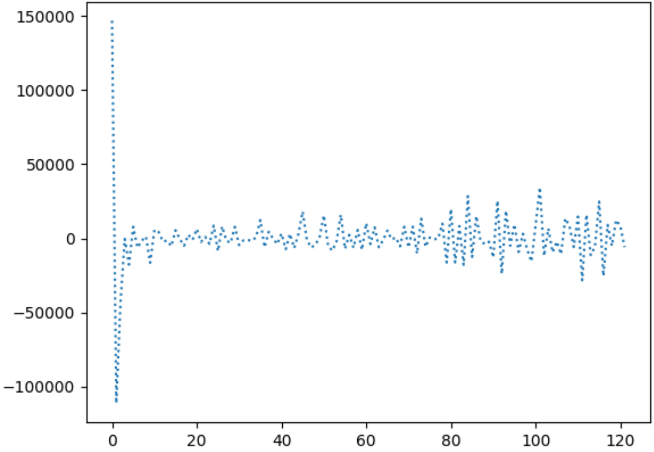
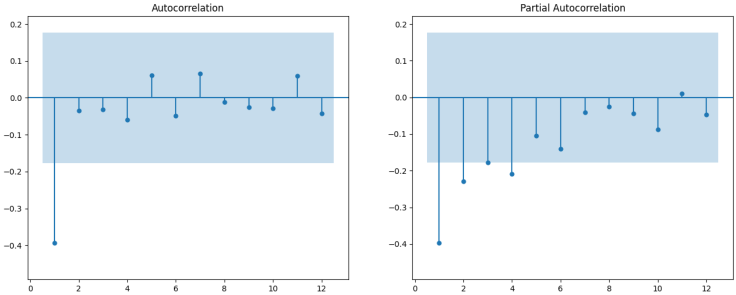
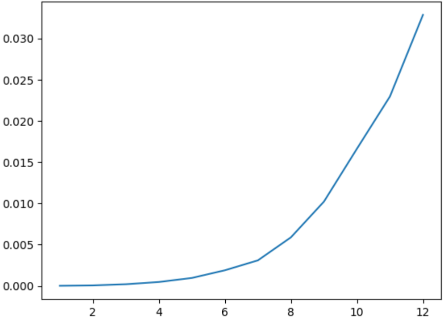


Figure: Residual for ARIMA model

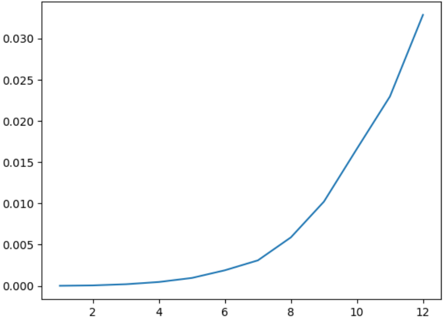
ARIMA(2,3,3): Ljung-box test



ARIMA(2,3,3): Ljung-box test



Ljung-box test p-value of ARIMA(2,3,3)



Ljung-box test p-value of GARCH(1,1)

According to the ljung-box test, we find that we reject the null hypothesis under the 95% significance level. Thus, we build a GARCH model with order (1,1) to modify the correlation of the residual.

ARIMA(2,3,3): prediction

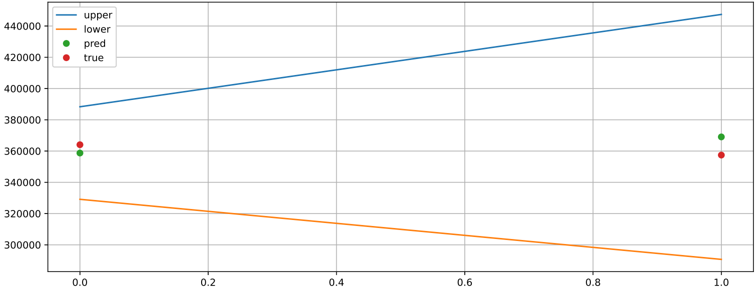


Figure: ARIMA(2,3,3) prediction and boundary

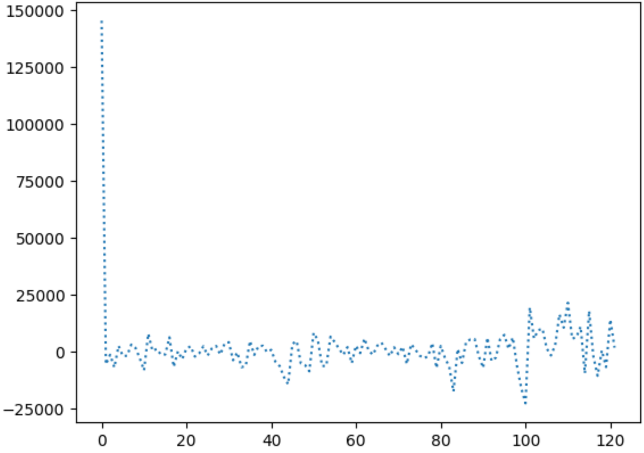
part1 model: SARIMA

Dep. Variable:	y	**No. Observations:**	122
Model:	SARIMAX(0, 1, 1)x(0, 0, 1, 4)	**Log Likelihood**	-1238.728
Date:	Thu, 08 Jun 2023	**AIC**	2485.457
Time:	08:59:53	**BIC**	2496.640
Sample:	0 - 122	**HQIC**	2489.999
Covariance Type:	opg		

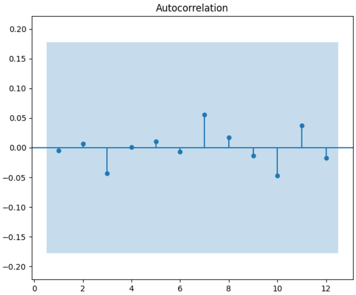
	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
intercept	1449.9972	671.475	2.159	0.031	133.930	2766.065
ma.L1	0.1181	0.045	2.626	0.009	0.030	0.206
ma.S.L4	-0.0737	0.072	-1.027	0.305	-0.214	0.067
sigma2	4.644e+07	0.008	5.59e+09	0.000	4.64e+07	4.64e+07

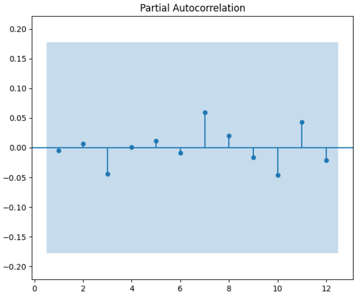
Ljung-Box (L1) (Q):	2.07	**Jarque-Bera (JB):**	14.83
Prob(Q):	0.15	**Prob(JB):**	0.00
Heteroskedasticity (H):	7.25	**Skew:**	0.06
Prob(H) (two-sided):	0.00	**Kurtosis:**	4.71

SARIMA(0,1,1): Residual

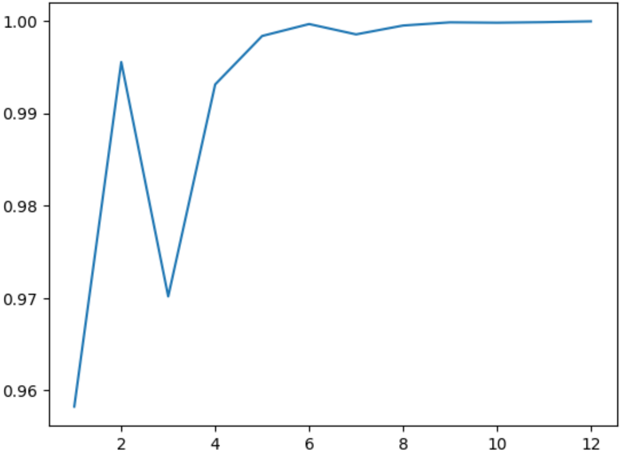


SARIMA(0,1,1): Residual



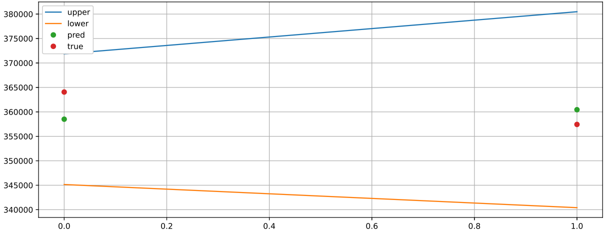


SARIMA(0,1,1): Residual



According to the ljung-box test, we find that we do not reject the null hypothesis under the 95% significance level.

SARIMA(0,1,1): prediction



Granger causality test: real part

	0	1	2	3	4	5	6	7	8	9
0	NaN	NaN	NaN	0.0014	NaN	NaN	NaN	NaN	NaN	0.0055
1	0.0029	NaN	0.0028	0.0046	NaN	0.0076	0.0034	0.0117	NaN	0.0003
2	0.0065	NaN	NaN	NaN	NaN	0.0065	0.0032	0.0027	NaN	0.0006
3	0.0000	0.0182	0.0017	NaN	NaN	0.0115	0.0001	0.0002	0.0006	0.0000
4	0.0001	NaN	0.0071	0.0010	NaN	NaN	0.0402	0.0055	NaN	0.0448
5	0.0010	0.0230	0.0033	0.0000	NaN	NaN	0.0065	0.0331	0.0004	0.0002
6	NaN	NaN	NaN	0.0000	0.0094	0.0053	NaN	NaN	0.0114	0.0000
7	0.0000	0.0338	0.0000	0.0000	0.0008	0.0011	0.0437	NaN	0.0089	0.0000
8	0.0048	0.0469	0.0079	0.0000	0.0205	0.0009	0.0100	NaN	NaN	0.0008
9	0.0053	0.0002	NaN	0.0000	NaN	0.0047	0.0002	0.0001	0.0089	NaN

Granger causality test: imaginary part

	0	1	2	3	4	5	6	7	8	9
0	NaN	0.0023	NaN	0.0014	NaN	0.0007	NaN	NaN	0.0024	NaN
1	0.0006	NaN	0.0324	0.0051	NaN	0.0056	0.0184	NaN	0.0027	NaN
2	0.0004	0.0075	NaN	0.0032	0.0105	0.0000	0.0006	NaN	0.0155	NaN
3	0.0001	0.0009	NaN	NaN	NaN	0.0000	0.0000	NaN	0.0263	NaN
4	0.0008	NaN	0.0050	0.0070	NaN	0.0006	NaN	0.0013	0.0007	NaN
5	0.0024	0.0025	0.0000	0.0029	NaN	NaN	0.0001	0.0099	NaN	NaN
6	0.0280	0.0011	0.0025	0.0002	0.0468	0.0326	NaN	0.0386	0.0077	NaN
7	0.0000	0.0089	0.0000	NaN	NaN	NaN	0.0014	NaN	NaN	NaN
8	0.0012	0.0065	0.0255	NaN	0.0013	NaN	0.0206	0.0000	NaN	NaN
9	0.0036	NaN	0.0227	0.0237	NaN	NaN	NaN	NaN	NaN	NaN

We exclude the ninth variable, which is close to zero and represents the symmetric axis, and include the rest in the VAR model for estimation.

part2 model: Estimation of VAR

$VAR(p)$ model for d-dimensional vector Y_t is given by:

$$Y_t = \alpha + \delta t + \Theta_1 Y_{t-1} + \cdots + \Theta_p T_{t-p} + \epsilon_t, \tag{3}$$

where $\alpha, \delta \in \mathbb{R}$, Θ_j is a $d \times d$ matrix and ϵ_t is white noise term belongs to $\mathbb{R}$.

part2 model: Forecasting of VAR

If we determine there do not exist the autocorrelation and cross-correlation in residuals of fitted model, then we can forecast the data as follows:

$$Y_{t+1} = \alpha + \delta(t+1) + \Theta_1 Y_t + \cdots + \Theta_p Y_{t-p+1}, \tag{4}$$

where $\alpha, \delta \in \mathbb{R}$ is a estimated vectors, Θ_j is a estimated $d \times d$ matrix and we ignore the white noise terms.

part2 model: VAR

```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Thu, 08, Jun, 2023
Time:                     11:00:26
--------------------------------------------------------------------
No. of Equations:         10.0000    BIC:                    121.942
Nobs:                     112.000    HQIC:                   107.374
Log likelihood:          -6035.12    FPE:                2.23386e+47
AIC:                      97.4270    Det(Omega_mle):     3.60944e+44
--------------------------------------------------------------------
```

After fitting VAR models, we get the max lag 10.

part3 model: VAR

```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Thu, 08, Jun, 2023
Time:                     11:05:18
--------------------------------------------------------------------
No. of Equations:         10.0000    BIC:                    43.9991
Nobs:                     121.000    HQIC:                   42.4897
Log likelihood:          -4115.09    FPE:                1.01605e+18
AIC:                      41.4574    Det(Omega_mle):     4.25629e+17
--------------------------------------------------------------------
```

After fitting VAR models, we get the max lag 10.

Durbin-Watson test for part 2 & part 3

After fitting VAR model, we determine the correlation among differnet time series data.

```
0 : 1.84115
1 : 2.16601
2 : 1.82301
3 : 2.10483
4 : 1.81346
5 : 1.51105
6 : 2.06541
7 : 1.54137
8 : 2.23077
9 : 2.01135
```

```
0 : 1.87956
1 : 2.23116
2 : 1.92861
3 : 2.09822
4 : 2.22694
5 : 2.02425
6 : 2.07579
7 : 1.93667
8 : 1.99370
9 : 2.42938
```

If the quantity is between 1.5 and 2.5, we claim that there is no significant auto-correlation among the residuals.

Table of Contents

Predictiton: ARIMA + VAR

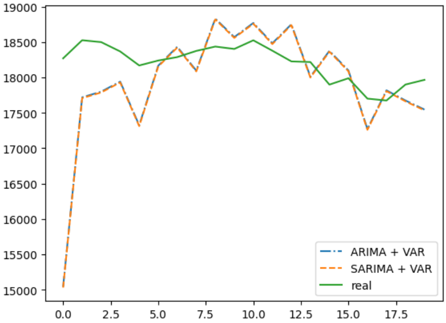


Figure: 1-month prediction result

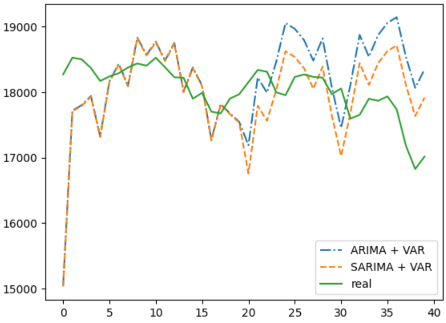


Figure: 2-month prediction result

Predictiton: ARIMA-GARCH(1,1) + VAR

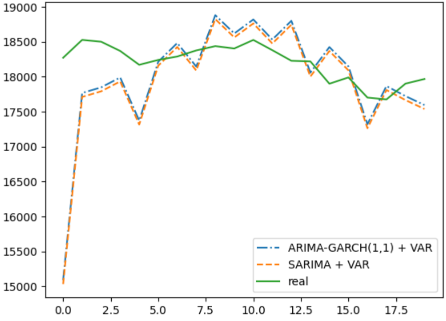


Figure: 1-month prediction result

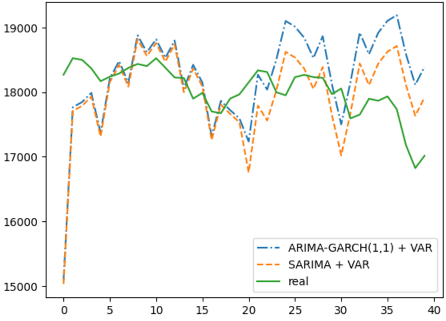


Figure: 2-month prediction result

Q & A

Q : What information in the time domain corresponds to image part in the frequency domain?

Q & A

A : In terms of algebraic representation in the discrete Fourier transform, we observe that each element of the frequency domain data is influenced by $x_1, ..., x_n$ in the time domain. The imaginary part of the frequency domain data originates from certain x_i values. Therefore, determining the corresponding relationship can be relatively complex. We assume that the relationship is a continuous function. Then, we speculate that a neural network can be used to attempt to discover this correspondence.

Q & A

Specifically, let neural network N_s.
output $\hat{P}$: randomly sampling multiple sets of time series data.
input X : the imaginary part of the data obtained after Fourier transforming these samples in the frequency domain.
structure : full connected layers.
loss function : $||\hat{P} - N_s(X)||$.

If the neural network converges during training, the trained network will represent the desired corresponding relationship we are seeking.